Technical Report: AIC-01-004, Naval Research Laboratory, Code 5515
4555 Overlook Avenue SW, Washington, DC, 20375-5537, June 12, 2001.

Semi-automatically Restructuring Navy Lessons: A Proposal for Feasibility Assessment and Progress Report

[1,2]Kalyan Moy Gupta, [1]David W. Aha, [1]Elaine Marsh, and [1]Tucker Maney

[1]Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Code 5515
4555 Overlook Avenue SW, Washington, DC, 20375-5537

[2]AES Division, ITT Industries,
2560 Huntington Avenue, Alexandria, VA, 22303-1410

Email: lastname@aic.nrl.navy.mil

Executive Summary

Lessons are semi-structured documents that record an organization's experience while conducting its operations. It is believed that reusing these lessons could improve the organization's decision-making and problem solving performance. However, existing lessons, such as those recorded by the Navy lessons learned system, lack adequate structure and content, which prevents their timely retrieval and effective reuse. In this report we propose that existing Navy lessons could potentially be restructured by applying information extraction techniques. This restructuring, we believe, will enable their retrieval using conversational case-based reasoning and thereby improve their chances of reuse and application. We provide an overview of the traditional information extraction (IE) task and outline our approach in adapting and applying it to the task of *navy lessons restructuring* (NLR). We argue that NLR is novel and complex in comparison to the traditional IE and warrants further investigation. We propose to examine the feasibility of applying IE to NLR. We conclude the report by summarizing the current status of our NLR efforts.

1. Restructuring Lessons: Motivation

Lessons are a type of knowledge artifact that *systematically* documents an organization's experience (failures and successes) in the context of its operations. The current *belief* is that applying lessons to affected processes can improve its performance. For example, lessons documenting the problematic factors encountered while executing previous naval training exercises could be used to anticipate and avoid similar problems in future exercises and operations.

Given this perception of anticipated benefits from reusing lessons, many organizations document their experiences at varying levels of structure, store them in electronic repositories, and provide a means to search them. We refer to this activity of collecting, organizing, storing, and disseminating lessons as Lessons Management (Aha & Gupta, 2001). Organizations engaged in lessons management include, among others, the US Navy (NLLS, 2001) and NASA. The most common lesson structure includes a set of fields with unstructured text as their values. The retrieval system primarily includes a keyword or full-text search. However, there is little but anecdotal evidence, principally from the champions of the system, about the extent of their reuse and the resulting benefits (Weber *et al.*, 2001). Among the reasons posited explaining the lack of use include the following:

1. *Poor quality and content of lessons.* The recorded experiences do not contain much information that complements and enhances the end user's knowledge.
2. *Less than adequate coverage of topics*. The recorded lessons do not adequately cover the topics on which the end users seek decision support (i.e., lesson bases are knowledge poor).
3. *Lack of adequate techniques and tools for lesson dissemination.* The simplistic representation of lessons and keyword retrieval techniques are inadequate for supporting complex problem solving.
4. *Lack of integration between lesson reuse and targeted organization process*. Lesson retrieval is not integrated with the decision support tools (Aha *et al*., 2001).
5. *Lack of institutional directives.* The organizations not adequately promoting and facilitating lesson reuse.

One of our goals in our project is to examine and increase the lesson quality and their coverage. Yet another is to significantly enrich the lesson representation to facilitate flexible, accurate, and timely application of lessons. While we can build lesson collection tools that support a richer structure and ensure that good quality lessons are authored in the future (e.g., LMTS [Aha & Gupta, 2001]), the issue of restructuring existing lessons remains. Clearly, efforts to manually restructure lessons will be prohibitive. Hence, we are investigating methods to semi-automatically perform this task. This task falls within the domain of information extraction.

2. Information Extraction (IE): Overview

IE is a computational technique for labeling and structuring semi-structured text into richer representations to enable a reasoning system make a wider range of inferences with greater accuracy than was possible prior to its restructuring and extraction.

Information extraction is an emerging field gaining in importance due to the recent explosion in the available semi-structured text and the need to exploit the information contained in them. Examples of extraction tasks range from extracting the succession events of senior business executives from raw financial newswire texts to the extraction of terrorist acts from the same (MUC-6, 1995). Extraction tasks typically include the following steps and resources (See Figure 1):

1. *Syntactic analysis.* This involves techniques ranging from part-of-speech tagging, sentence segmentation, and morphological analysis to complex sentence structure analysis. Depending on the nature of the input corpus, shallow parsing techniques can be successfully used. This step may be skipped for corpi with telegraphic (i.e., cryptic or highly abbreviated) expressions such as equipment maintenance logs.
2. *Semantic analysis*. This involves assigning semantic classes to the various text elements (i.e., words, phrases) by using a combination of lexical resources and semantic tagging techniques. The process of assigning semantic classes can be quite complex and often involves using machine learning techniques. In its simplistic form,

semantic tagging involves looking up a class in the lexicon for a given word-form or text entity. Examples include identification of persons, organizations, and places. The sophistication required of a semantic tagging technique increases with the subject matter scope of input documents, which increases the potential number of senses for word forms (i.e., lexical ambiguity). The complexity is compounded by the codependency of syntactic and semantic processing techniques.

Typically, it is assumed that the lexical resources for semantic tagging are available. For very narrow and well-defined extraction tasks such lexical resources are manually developed. Lexical resources may include semantic patterns or relations for use in semantic analysis and extraction. These are relational structures composed of a mixture of word, syntactic (i.e., part-of-speech tags), and semantic (i.e., sense information) elements. Clearly, the patterns composed entirely of semantic elements are more general than those that only use syntactic or word form elements. For very narrow topic such as terrorist acts or rental advertisement, the semantic patterns can be manually coded. However, manual coding of patterns is error-prone, effort intensive, and requires skilled knowledge engineers. Consequently, methods for learning patterns have been suggested and used successfully (Soderland, 1999). Regardless, of how the semantic patterns are derived they are central to the success of the extraction task.

3. *Extraction*: This is the process of instantiating and populating a target case frame structure with the elements from the source text. In its simple form, the target case frame could be a relational database record with a set of fields. For example, a rental advertisement may include a record with fields such as location, price, and number of bedrooms. Usually, complex target structures comprise case frames with multiple slots that assume a restricted set of values. Typically, extraction involves placing the semantically tagged text entities into their respective slots in the case frames.

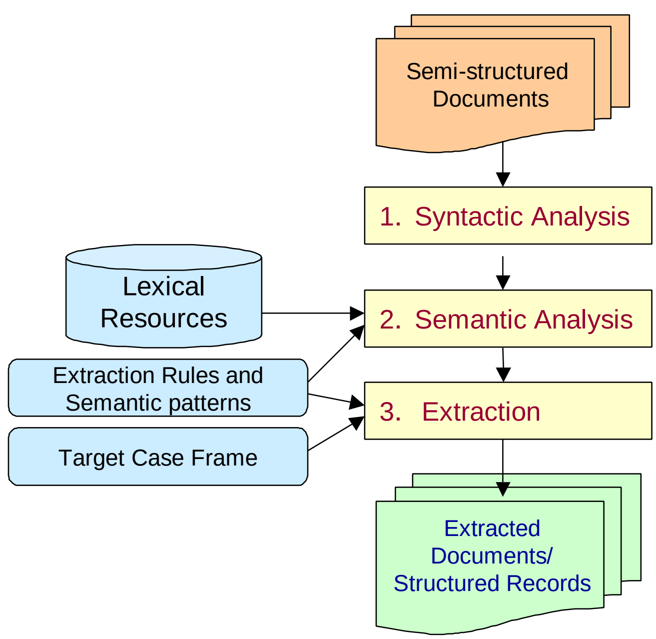

Figure 1 Overview of Information Extraction

3. Semi-automatically Restructuring Navy Lessons

The US Navy presently collects and disseminates lessons through the Navy Lessons Learned System (NLLS). It suffers from the same deficiencies as those discussed in Section 1, including limited representation and dissemination tools. Since there were approximately 40,000 unclassified recorded entries in NLLS repository in the beginning of the year 2001, which continues to grow at the rate of over 2000 per year, we conjectured that restructuring these lessons could substantially benefit Navy end users.

To make a lesson useful and applicable to target decision processes its representation should capture the domain complexity and the detail needed for reasoning and decision support. Our goal in this project is to *restructure the existing semi-structured Navy lessons* (NLLS, 2001) into a *target case frame structure* and enable its retrieval in a flexible and conversational way. In particular, we will focus on approximately 12,000 Navy active and inactive lessons contained in NLLS. In the remaining report, we will refer to this goal as ***Navy Lessons Restructuring*** (NLR).

For this project, we will use the LMTS lesson representation as our target case frame structure developed earlier (See www.aic.nrl.navy.mil/lmts). The LMTS lesson structure was developed for collecting NASA Design for Safety program lessons. The principal characteristics of this representation include task and user context, content and context generalization using taxonomies, and implementation of the problem-solving framework comprising observation, analysis and recommendation. This representation extends and enriches the existing NLLS lesson structure and would support taxonomic conversational retrieval of lessons (Gupta, 2001).

The LMTS representation includes slots such as lesson context, lesson observations or conditions of applicability, lesson learned, and corrective actions. In LMTS, the slots take values that are nodes in feature taxonomies. As a part of lesson restructuring process, the feature taxonomies must be extracted and developed. Therefore, the NLR task comprises two parts: (1) extracting feature taxonomies, and (2) populating one LMTS case frame structure for each existing NLLS lesson.

While we expect to apply many of the traditional IE approach to the NLR task, we anticipate extensions of the same and the development of new techniques. We believe that NLR is harder than the traditional IE tasks due to the following reasons:

1. *The target case frame complexity:* LMTS lesson structure has many slots each of them is potentially multi-valued. The values themselves need to be extracted and could involve sub-case frame structures. For example, observations or conditions of applicability are target structures themselves and so is a recommended action. In other words, this makes NLR a *compound information extraction task*.
2. *Large scope of the subject matter*: Navy lessons pertain to complex and diverse organization processes ranging from planning naval exercises to scheduling training for marine personnel. The scope of subject matter is an order of magnitude larger than the scope of traditional extraction task such as terrorist event extraction or business executives succession event structuring. A more limited scope equivalent to the terrorist act extraction would be, for example, restructuring Navy lessons pertaining to subject of “Train Forces and Personnel.”

Given this, the performance of traditional IE techniques on NLR needs to be investigated.

4. Examining NLR Feasibility

In the beginning of this project, our approach was to develop tools for automating all the NLR steps (See Figure 2). However, our preliminary assessment of the NLR task leads us to believe that it is novel and potentially more complex than the traditional IE tasks presented in Section 2. Furthermore, we observed that the development of tools for automatic acquisition of lexical resources was complex and tools for syntactic and semantic processing were limited in their performance. Given the complexity and the novelty of NLR, we propose a revised approach to expeditiously assess the feasibility of the extraction step. We will manually prepare the lexical resources and augment the syntactic and semantic processing steps with manual inputs. This will allow us to examine extraction under assumptions of perfect syntactic and semantic preprocessing and enable us to gain insight into the impact of preprocessing on the extraction task. If the extraction performance is acceptable we can return to the task of developing tools for automatic acquisition of lexical resources and improving the preprocessing steps.

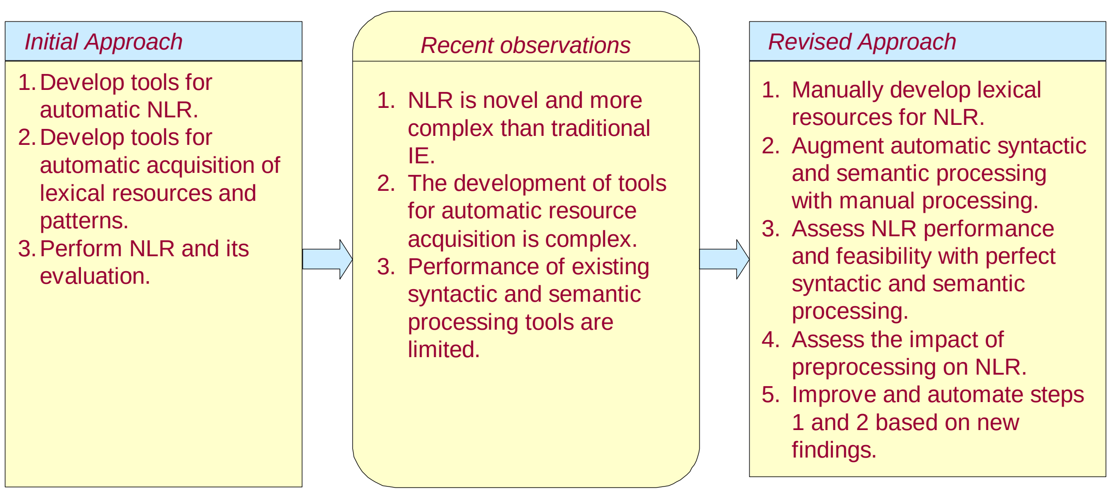

Figure 2. Rationale for our proposed approach to NLR evaluation

The revised proposal to examine the feasibility of NLR is as follows:

1. *Select a narrow target domain*: Select a small cohesive set of lessons on one or more topics (e.g., Train Forces and Personnel). We have prepared the distribution of 12317 navy active and inactive lessons by their Universal Naval Task List (NTA) codes.

2. *Decompose NLR task into a collection of simpler extraction sub-tasks*: Decompose the extraction task into extraction of sub case-frame structures for each slot in the LMTS case frame. For example, prepare and extract sub-case frame structure for tasks, observations, and recommendation slots in the LMTS case frame structure.

3. *Develop extraction performance measures and expectations:* Typically extraction performance is measured by recall and precision. We will extend and apply recall

and precision measures to the tasks identified in step 2 and establish performance expectations by manually performing various NLR steps (See step 4).

4. *Assess the feasibility of manual NLR*: Assess the feasibility and the complexity of the NLR task and its sub-tasks by performing a manual restructuring exercise on a few sample lessons.

5. *Assess the feasibility of semi-automatic NLR*: Although the steps for IE remain the same for various tasks, they can vary considerably in the degree of sophistication. To expeditiously test the feasibility of *NLR*, we will perform the following (See Figure 2):

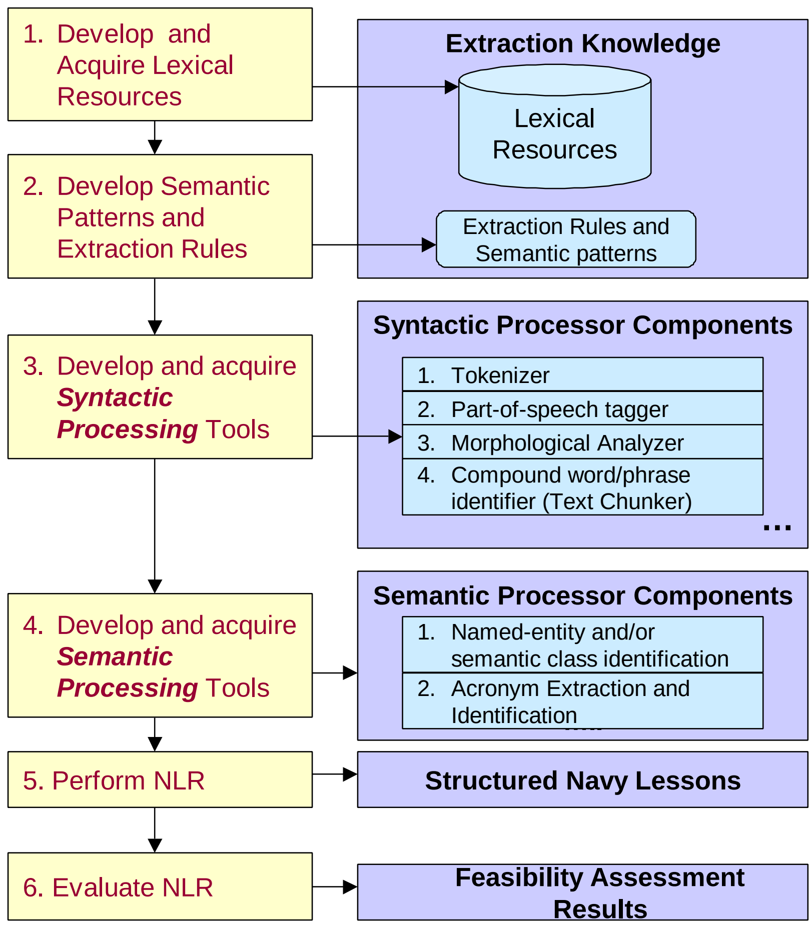

Figure 3. Assessing the feasibility of semi-automatic NLR

a. *Manually develop lexical resources*: We will manually develop the necessary lexicons for the selected subject matter and the manually tag lessons for their semantic components where needed.
b. *Manually develop semantic patterns:* Using the elements in the lexicons we will develop the semantic patterns and extraction rules relevant to the NLR task.
c. *Develop, acquire, and adapt syntactic processing software components:* We will acquire and adapt the syntactic processing tools such as part of

speech taggers, morphological analyzers, text chunking and phrase identification software.

d. *Develop, acquire and adapt semantic processing software components*: Any available semantic tagging software will be acquired and extended.

e. *Evaluate NLR performance on the selected test set*: We will execute the extraction routines comprising syntactic processing, semantic processing, and extraction on the selected Navy lessons and estimate performance by comparison to our manual extraction performance.

5. Status of NLR task

The work on NLR was started 8 February 2001. A framework of necessary tools needed to perform the lesson extraction task was proposed. According to the framework, the development of syntactic analysis tools and preparation of lexical resources were initiated. The following tasks were completed (See Figure 2 for comparison):

1. *Development of lexical resources:* Extensive literature survey was undertaken on the techniques for acquiring lexical resources and learning extraction pattern rules. WordNet was acquired and transformed into MS Access format for further use. A Navy training and planning task specific lexicon is being manually crafted.

2. Development of semantic patterns and resources: The semantic patterns will be developed after the development of domain specific lexicons are completed.

3. *Development and acquisition of syntactic processing tools*: The tools developed and adapted to date include:

    a. Adaptation of a Java part-of-speech tagger called QTAG (See Qtag, 2001)

    b. A simple morphological analyzer to obtain root form of various words using WordNet.

    c. We attempted the task of noun-phrase detection and more generally text-chunking. We investigated memory-based text chunking by implementing Stanfill Waltz (1986) memory-based reasoning (MBR) in Java. The routine was tested on CoNLL 2000 text chunking task test set (CoNLL, 2000). The results of our experiment were comparable to those obtained by Veenstra, and van den Bosch (2000). It produced higher precision (93.7 vs. 91.2 %) on noun tags compared to Veenstra and van den Bosch. This was due to the reason that our representation included chunk tags in addition to word forms and part of speech tags. The results have not been published yet.

    d. Sentence structure identification using link-parser: A sentence structure parser called Link Parser (Temperly *et al.*, 2001) has been adapted for performing syntactic analysis.

3. *Development and acquisition of semantic processing tools*: The following efforts were directed toward the acquisition of semantic processing tools:

    a. Sense tagging and semantic disambiguation: Literature survey on semantic disambiguation, entity co-reference resolution, and development of semantic primitives for a sub-domain language was undertaken.

b. Acronym extraction and identification development: Since the lessons contain numerous acronyms, it was felt that they must be resolved with their expansion in the text. A literature search was conducted and a heuristic acronym extractor was implemented. The extractor was run on 4200 Navy lessons to extract over 700 unique acronyms and their corresponding expansions. Precision of over 80% was achieved with this implementation. The precision was computed by counting the incorrect acronym extractions over the complete extracted set.

4. The navy lessons text data has been processed into MS Access format for easy selection of test set.

6. Conclusion and Future work

Lesson management is gaining importance in large organizations that involves the collection, organization and dissemination of problem-solving and decision-making experience. While new methods and tools for improving lesson management are being proposed, we see the problem of transforming existing lessons to work with the new methods and processes an important one. To this end, in this report, we propose to restructure the existing active and inactive Navy lessons contained in the NILLS repository by employing IE techniques. Based on our preliminary investigation, we concluded that NLR is more complex than traditional IE and that it warrants a detailed evaluation. We outline our strategy for evaluating the feasibility of NLR.

In the near future, we will complete the NLR evaluation and report our findings. We plan to extend this research and investigate the applicability of our approach to other lesson management systems. Over the long term we intend to investigate the more complex task of lesson identification and extraction from unstructured documents and emails.

Acknowledgement

This research was funded by the Design For Safety Program of the NASA Ames Research Center, Naval Research Laboratory, and the Office of Naval Research.

References

Aha, D.W., and Gupta, K.M., (2001). Lesson Management Tool Suite Project Report, *Technical Report AIC-01-001*, Naval Research Laboratory, Navy Center for Applied Research in AI, Washington, DC.

Aha, D.W., Breslow, L.A., & Muñoz-Avila, H. (2001). Conversational Case-Based Reasoning. *Applied Intelligence*, *14*, 9-32.

Aha, D.W., Weber, R., Muñoz, H., Breslow, L.A. & Gupta, K. (2001). Bridging the Lesson Distribution Gap. *To appear in the proceedings of IJCAI'01*.

CoNLL (2000), *CoNLL text chunking task*, http://lcg-www.uia.ac.be/conll2000/chunking/

Gupta, K.M.,(2001). Taxonomic Conversational Case-Based Reasoning, to Appear In *ICCBR-2001*.

Miller, G.A., Beckwith, R., Christiane, F., Gross, D., and Miller, K., (1993) Introduction to WordNet.

MUC-6, (1995). *Proceedings of the Sixth Message Understanding Conference*, Morgan Kaufman, San Francisco, CA.

NLLS (2001), Navy Lessons Learned System, http://www.nwdc.navy.mil/Command/NavyLessonsLearned/nllsoverview.asp

Qtag (2001), A Portable Part-of-Speech Tagger, http://www-clg.bham.ac.uk/QTAG/

Stanfill C., and Waltz, D. (1986). Toward Memory-Based Reasoning, *Communications of the ACM*, 29(12), pp. 1213-1228.

Soderland, S., (1999), Learning Information Extraction Rules for Semi-structured and Free Text, *Machine Learning*, 34(1-3): pp. 233-272.

Temperly,D., Sleator, D., & Laferty, J., (2001) Link Grammar, www.link.cs.cmu.edu/link

Veenstra, J., and van den Bosch, A., Single-Classifier Memory-based Phrase Chunking, *Proceedings of the CoNLL-2000 and LLL-2000*, Lisbon, Portugal.

Weber, R., Aha, D.W., Becerra-Fernandez, I., (2001). Intelligent Lessons Learned Systems, *Expert Systems with Applications*, 17(1), pp. 17-34.

Zernik, U., (1991). *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, (Ed.) U. Zernik, Lawrence Earlbaum Associates, Hillsdale, NJ.